#5 PCT09

ENTERED

RAW SEQUENCE LISTING   DATE: 02/27/2001
PATENT APPLICATION: US/09/762,926   TIME: 14:54:35

Input Set : A:\seqlist.txt
Output Set: N:\CRF3\02272001\I762926.raw

```
     4 <110> APPLICANT: Thonnard, Joelle
     6 <120> TITLE OF INVENTION: Novel Compounds
     9 <130> FILE REFERENCE: BM45330
C--> 11 <140> CURRENT APPLICATION NUMBER: US/09/762,926
C--> 11 <141> CURRENT FILING DATE: 2001-02-14
    11 <150> PRIOR APPLICATION NUMBER: PCT/EP99/05989
    12 <151> PRIOR FILING DATE: 1999-08-13
    14 <150> PRIOR APPLICATION NUMBER: GB 9818004.5
    15 <151> PRIOR FILING DATE: 1998-08-18
    17 <160> NUMBER OF SEQ ID NOS: 9
    19 <170> SOFTWARE: FastSEQ for Windows Version 4.0
    21 <210> SEQ ID NO: 1
    22 <211> LENGTH: 2769
    23 <212> TYPE: DNA
    24 <213> ORGANISM: Bacteria
    26 <400> SEQUENCE: 1
    27 atgagatctt ctttccggtt gaagccgatt tgtttttatc ttatgggtgt tatgctatat      60
    28 catcatagtt atgccgaaga tgcagggcgc gcgggcagcg aggcgcagat acaggttttg     120
    29 gaagatgtgc acgtcaaggc gaagcgcgta ccgaaagaca aaaaagtgtt taccgatgcg     180
    30 cgtgccgtat cgacccgtca ggatatattc aaatccagcg aaaaacctcga caacatcgta     240
    31 cgcagcatcc ccggtgcgtt tacacagcaa gataaaagct cgggcattgt gtctttgaat     300
    32 attcgcggcg acagcgggtt cgggcgggtc aatacgatgg tggacggcat cacgcagacc     360
    33 ttttattcga cttctaccga tgcgggcagg gcaggcggtt catctcaatt cggtgcatct     420
    34 gtcgacagca attttattgc cggactggat gtcgtcaaag gcagcttcag cggctcggca     480
    35 ggcatcaaca gccttgccgg ttcggcgaat ctgcggactt taggcgtgga tgacgtcgtt     540
    36 cagggcaata atacctacgg cctgctgcta aaaggtctga ccggcaccaa ttcaaccaaa     600
    37 ggtaatgcga tggcggcgat aggtgcgcgc aaatggctgg aaagcggagc atctgtcggt     660
    38 gtgcttttacg ggcacagcag gcgcacgtgg gcgcaaaatt accgcgtggg cggcggcggg     720
    39 cagcacatcg gaaattttgg cgcggaatat ctggaacggc gcaaacagcg atattttgta     780
    40 caagaaggcg ggttgaaatt caattccaac agcggaaaat gggagcggga tttccaaagg     840
    41 ccgtactgga aaaccaagtg gtatcaaaaa tacaatgacc cccaagaact gcaaaaatac     900
    42 atcgaaggtc atgacaaaag ctggcgggaa aacctggcgc cgcaatacga catcaccccc     960
    43 atcgatccgt ccagcctgaa gcagcagtcg gcaggcaatc tgtttaaatt ggaatacgac    1020
    44 ggcgtattca ataaatacac ggcgcaattt cgcgatttaa acaccaaaat cggcagccgc    1080
    45 aaaatcatca accgcaatta tcaattcaat tacggtttat ctttaaactc atatgccaac    1140
    46 ctcaatctga ccgcagccta caattcgggc aggcagaaat atccgaaagg gtcgaagttt    1200
    47 acaggctggg ggcttttaaa agattttgaa acctacaaca acgcgaaaat cctcgacctc    1260
    48 aacaacaccg ccaccttccg gctgccccgc gaaaccgagt tgcaaaccac tttgggcttc    1320
    49 aattatttcc acaacgaata cggcaaaaac cgctttcctg aagaattggg gctgtttttc    1380
    50 gacggtccgg atcaggacaa cgggctttat tcctatttgg ggcggtttaa gggcgataaa    1440
    51 gggctgctgc cccaaaaatc aaccatcgtc caaccggccg gcagccaata tttcaacacg    1500
    52 ttctacttcg atgccgcgct caaaaaagac atttaccgct taaactacag caccaatacc    1560
    53 gtcggctacc gtttcggcgg cgaatatacg ggctattacg gctcggatga cgaatttaag    1620
    54 cgggcattcg gagaaaactc gccgacatac aagaaacatt gcaaccagag ctgcggaatt    1680
    55 tatgaacccg tattgaaaaa atacggcaaa aagcgcgcca acaaccattc ggtcagcatt    1740
    56 agtgcggact tcggcgatta tttcatgccg ttcgccagct attcgcgcac acaccgtatg    1800
    57 cccaacatcc aagaaatgta tttttcccaa atcggcgact ccggcgttca caccgcctta    1860
```

```
58 aaaccagagc gcgcaaacac ttggcaattt ggcttcaata cctataaaaa aggattgtta   1920
59 aaacaagatg atacattagg attaaaactg gtcggctacc gcagccgcat cgacaactac   1980
60 atccacaacg tttacgggaa atggtgggat ttgaacggga atattccgag ctgggtcagc   2040
61 agcaccgggc ttgcctacac catccaacac cgcaatttca aagacaaagt acacaaacac   2100
62 ggttttgagt tggagctgaa ttacgattat gggcgttttt tcaccaacct ttcttacgcc   2160
63 tatcaaaaaa gcacgcaacc gaccaacttc agcgatgcga gcgaatcgcc caacaatgcg   2220
64 tccaaagaag accaactcaa acaaggttat gggttgagca gggtttccgc cctgccgcga   2280
65 gattacggac gtttggaagt cggtacgcgc tggttgggca acaaactgac tttgggcggc   2340
66 gcgatgcgct atttcggcaa gagcatccgc gcgacggctg aagaacgcta tatcgacggc   2400
67 accaacgggg gaaataccag caatgtccgg caactgggca agcgttccat caaacaaacc   2460
68 gaaacccttg cccgccagcc tttgattttt gatttttacg ccgcttacga gccgaagaaa   2520
69 aaccttattt tccgcgccga agtcaaaaat ctgttcgaca ggcgttatat cgatccgctc   2580
70 gatgcgggca atgatgcggc aacgcagcgt tattacagtt cgttcgaccc gaaagacaag   2640
71 gacgaagaag taacgtgtaa tgctgataaa acgttgtgca acggcaaata cggcggcaca.  2700
72 agcaaaagcg tattgaccaa ttttgcacgc ggacgcacct ttttgataac gatgagctac   2760
73 aagttttaa                                                           2769
75 <210> SEQ ID NO: 2
76 <211> LENGTH: 922
77 <212> TYPE: PRT
78 <213> ORGANISM: Bacteria
80 <400> SEQUENCE: 2
81 Met Arg Ser Ser Phe Arg Leu Lys Pro Ile Cys Phe Tyr Leu Met Gly
82  1               5                  10                  15
83 Val Met Leu Tyr His His Ser Tyr Ala Glu Asp Ala Gly Arg Ala Gly
84              20                  25                  30
85 Ser Glu Ala Gln Ile Gln Val Leu Glu Asp Val His Val Lys Ala Lys
86          35                  40                  45
87 Arg Val Pro Lys Asp Lys Lys Val Phe Thr Asp Ala Arg Ala Val Ser
88      50                  55                  60
89 Thr Arg Gln Asp Ile Phe Lys Ser Ser Glu Asn Leu Asp Asn Ile Val
90 65                  70                  75                  80
91 Arg Ser Ile Pro Gly Ala Phe Thr Gln Gln Asp Lys Ser Ser Gly Ile
92                  85                  90                  95
93 Val Ser Leu Asn Ile Arg Gly Asp Ser Gly Phe Gly Arg Val Asn Thr
94              100                 105                 110
95 Met Val Asp Gly Ile Thr Gln Thr Phe Tyr Ser Thr Ser Thr Asp Ala
96              115                 120                 125
97 Gly Arg Ala Gly Gly Ser Ser Gln Phe Gly Ala Ser Val Asp Ser Asn
98          130                 135                 140
99 Phe Ile Ala Gly Leu Asp Val Val Lys Gly Ser Phe Ser Gly Ser Ala
100 145                 150                 155                 160
101 Gly Ile Asn Ser Leu Ala Gly Ser Ala Asn Leu Arg Thr Leu Gly Val
102              165                 170                 175
103 Asp Asp Val Val Gln Gly Asn Asn Thr Tyr Gly Leu Leu Leu Lys Gly
104              180                 185                 190
105 Leu Thr Gly Thr Asn Ser Thr Lys Gly Asn Ala Met Ala Ala Ile Gly
106          195                 200                 205
107 Ala Arg Lys Trp Leu Glu Ser Gly Ala Ser Val Gly Val Leu Tyr Gly
108      210                 215                 220
```

```
109 His Ser Arg Arg Thr Trp Ala Gln Asn Tyr Arg Val Gly Gly Gly Gly
110 225                 230                 235                 240
111 Gln His Ile Gly Asn Phe Gly Ala Glu Tyr Leu Glu Arg Arg Lys Gln
112             245                 250                 255
113 Arg Tyr Phe Val Gln Glu Gly Gly Leu Lys Phe Asn Ser Asn Ser Gly
114             260                 265                 270
115 Lys Trp Glu Arg Asp Phe Gln Arg Pro Tyr Trp Lys Thr Lys Trp Tyr
116     275                 280                 285
117 Gln Lys Tyr Asn Asp Pro Gln Glu Leu Gln Lys Tyr Ile Glu Gly His
118     290                 295                 300
119 Asp Lys Ser Trp Arg Glu Asn Leu Ala Pro Gln Tyr Asp Ile Thr Pro
120 305                 310                 315                 320
121 Ile Asp Pro Ser Ser Leu Lys Gln Gln Ser Ala Gly Asn Leu Phe Lys
122             325                 330                 335
123 Leu Glu Tyr Asp Gly Val Phe Asn Lys Tyr Thr Ala Gln Phe Arg Asp
124             340                 345                 350
125 Leu Asn Thr Lys Ile Gly Ser Arg Lys Ile Ile Asn Arg Asn Tyr Gln
126             355                 360                 365
127 Phe Asn Tyr Gly Leu Ser Leu Asn Ser Tyr Ala Asn Leu Asn Leu Thr
128     370                 375                 380
129 Ala Ala Tyr Asn Ser Gly Arg Gln Lys Tyr Pro Lys Gly Ser Lys Phe
130 385                 390                 395                 400
131 Thr Gly Trp Gly Leu Leu Lys Asp Phe Glu Thr Tyr Asn Asn Ala Lys
132             405                 410                 415
133 Ile Leu Asp Leu Asn Asn Thr Ala Thr Phe Arg Leu Pro Arg Glu Thr
134             420                 425                 430
135 Glu Leu Gln Thr Thr Leu Gly Phe Asn Tyr Phe His Asn Glu Tyr Gly
136         435                 440                 445
137 Lys Asn Arg Phe Pro Glu Glu Leu Gly Leu Phe Phe Asp Gly Pro Asp
138     450                 455                 460
139 Gln Asp Asn Gly Leu Tyr Ser Tyr Leu Gly Arg Phe Lys Gly Asp Lys
140 465                 470                 475                 480
141 Gly Leu Leu Pro Gln Lys Ser Thr Ile Val Gln Pro Ala Gly Ser Gln
142             485                 490                 495
143 Tyr Phe Asn Thr Phe Tyr Phe Asp Ala Ala Leu Lys Lys Asp Ile Tyr
144             500                 505                 510
145 Arg Leu Asn Tyr Ser Thr Asn Thr Val Gly Tyr Arg Phe Gly Gly Glu
146         515                 520                 525
147 Tyr Thr Gly Tyr Tyr Gly Ser Asp Asp Glu Phe Lys Arg Ala Phe Gly
148     530                 535                 540
149 Glu Asn Ser Pro Thr Tyr Lys Lys His Cys Asn Gln Ser Cys Gly Ile
150 545                 550                 555                 560
151 Tyr Glu Pro Val Leu Lys Lys Tyr Gly Lys Lys Arg Ala Asn Asn His
152             565                 570                 575
153 Ser Val Ser Ile Ser Ala Asp Phe Gly Asp Tyr Phe Met Pro Phe Ala
154             580                 585                 590
155 Ser Tyr Ser Arg Thr His Arg Met Pro Asn Ile Gln Glu Met Tyr Phe
156     595                 600                 605
157 Ser Gln Ile Gly Asp Ser Gly Val His Thr Ala Leu Lys Pro Glu Arg
```

```
158      610                  615                  620
159 Ala Asn Thr Trp Gln Phe Gly Phe Asn Thr Tyr Lys Lys Gly Leu Leu
160 625                  630                  635                  640
161 Lys Gln Asp Asp Thr Leu Gly Leu Lys Leu Val Gly Tyr Arg Ser Arg
162                  645                  650                  655
163 Ile Asp Asn Tyr Ile His Asn Val Tyr Gly Lys Trp Trp Asp Leu Asn
164                  660                  665                  670
165 Gly Asn Ile Pro Ser Trp Val Ser Ser Thr Gly Leu Ala Tyr Thr Ile
166              675                  680                  685
167 Gln His Arg Asn Phe Lys Asp Lys Val His Lys His Gly Phe Glu Leu
168          690                  695                  700
169 Glu Leu Asn Tyr Asp Tyr Gly Arg Phe Phe Thr Asn Leu Ser Tyr Ala
170 705                  710                  715                  720
171 Tyr Gln Lys Ser Thr Gln Pro Thr Asn Phe Ser Asp Ala Ser Glu Ser
172                  725                  730                  735
173 Pro Asn Asn Ala Ser Lys Glu Asp Gln Leu Lys Gln Gly Tyr Gly Leu
174              740                  745                  750
175 Ser Arg Val Ser Ala Leu Pro Arg Asp Tyr Gly Arg Leu Glu Val Gly
176          755                  760                  765
177 Thr Arg Trp Leu Gly Asn Lys Leu Thr Leu Gly Gly Ala Met Arg Tyr
178          770                  775                  780
179 Phe Gly Lys Ser Ile Arg Ala Thr Ala Glu Glu Arg Tyr Ile Asp Gly
180 785                  790                  795                  800
181 Thr Asn Gly Gly Asn Thr Ser Asn Val Arg Gln Leu Gly Lys Arg Ser
182          .        805                  810                  815
183 Ile Lys Gln Thr Glu Thr Leu Ala Arg Gln Pro Leu Ile Phe Asp Phe
184              820                  825                  830          .
185 Tyr Ala Ala Tyr Glu Pro Lys Lys Asn Leu Ile Phe Arg Ala Glu Val
186          835                  840                  845
187 Lys Asn Leu Phe Asp Arg Arg Tyr Ile Asp Pro Leu Asp Ala Gly Asn
188      850                  855                  860
189 Asp Ala Ala Thr Gln Arg Tyr Tyr Ser Ser Phe Asp Pro Lys Asp Lys
190 865                  870                  875                  880
191 Asp Glu Glu Val Thr Cys Asn Ala Asp Lys Thr Leu Cys Asn Gly Lys
192              885                  890                  895
193 Tyr Gly Gly Thr Ser Lys Ser Val Leu Thr Asn Phe Ala Arg Gly Arg
194          900                  905                  910
195 Thr Phe Leu Ile Thr Met Ser Tyr Lys Phe
196          915                  920
198 <210> SEQ ID NO: 3
199 <211> LENGTH: 2769
200 <212> TYPE: DNA
201 <213> ORGANISM: Bacteria
203 <400> SEQUENCE: 3
204 atgagatctt ctttccggtt gaagccgatt tgtttttatc ttatgggtgt tatgctatat      60
205 catcatagtt atgccgaaga tgcaggycqc gcyggcagcg aggcgcagat acaggttttg     120
206 gaagatgtgc acgtcaaggc gaagcgcgta ccgaaagaca aaaaagtgtt taccgatgcg     180
207 cgtgccgtat cgacccgtca ggatatattc aaatccagcg aaaacctcga caacatcgta     240
208 cgcagcatcc ccggtgcgtt tacacagcaa gataaaagct cgggcattgt gtctttgaat     300
```

RAW SEQUENCE LISTING   DATE: 02/27/2001
PATENT APPLICATION: US/09/762,926   TIME: 14:54:35

Input Set : A:\seqlist.txt
Output Set: N:\CRF3\02272001\I762926.raw

```
209 attcgcggcg acagcgggtt cgggcgggtc aatacgatgg tggacggcat cacgcagacc       360
210 ttttattcga cttctaccga tgcgggcagg gcaggcggtt catctcaatt cggtgcatct       420
211 gtcgacagca attttattgc cggactggat gtcgtcaaag gcagcttcag cggctcggca       480
212 ggcatcaaca gccttgccgg ttcggcgaat ctgcggactt taggcgtgga tgacgtcgtt       540
213 caggcaata  atacctacgg cctgctgcta aaaggtctga ccggcaccaa ttcaaccaaa       600
214 ggtaatgcga tggcggcgat aggtgcgcgc aaatggctgg aaagcggagc atctgtcggt       660
215 gtgctttacg ggcacagcag gcgcagcgtg gcgcaaaatt accgcgtggg cggcggcggg       720
216 cagcacatcg gaaattttgg cgcggaatat ctggaacggc gcaagcagcg atattttgta       780
217 caagaaggcg ggttgaaatt caattccaac agcggaaaat gggagcggga tttccaaagg       840
218 ccgtactgga aaaccaagtg gtatcaaaaa tacaatgacc cccaagaact gcaaaaatac       900
219 atcgaaggtc atgacaaaag ctggcgggaa aacctgcgc  cgcaatacga catcacccc       960
220 atcgatccgt ccagcctgaa gcagcagtcg gcaggcaatc tgtttaaatt ggaatacgac      1020
221 ggcgtattca ataaatacac ggcgcaattt cgcgatttaa acaccaaaat cggcagccgc      1080
222 aaaatcatca accgcaatta tcaattcaat tacggtttat cttttaaactc atatgccaac      1140
223 ctcaatctga ccgcagccta caattcgggc aggcagaaat atccgaaagg gtcgaagttt      1200
224 acaggctggg ggcttttaaa agattttgaa acctacaaca acgcgaaaat cctcgacctc      1260
225 aacaacaccg ccaccttccg gctgccccgc gaaaccgagt gcaaaccac  tttgggcttc      1320
226 aattattcc  acaacgaata cggcaaaaac cgctttcctg aagaattggg gctgtttttc      1380
227 gacggtccgg atcaggacaa cgggctttat tcctatttgg ggcggtttaa gggcgataaa      1440
228 gggctgctgc cccaaaaatc aaccatcgtc caaccggccg gcagccaata tttcaacacg      1500
229 ttctacttcg atgccgcgct caaaaaagac atttaccgct taaactacag caccaatacc      1560
230 gtcggctacc gtttcggcgg cgaatatacg ggctattacg gctcggatga cgaatttaag      1620
231 cggcattcg  gagaaaactc gccgacatac aagaaacatt gcaaccagag ctgcggaatt      1680
232 tatgaacccg tattgaaaaa atacggcaaa aagcgcgcca acaaccattc ggtcagcatt      1740
233 agtgcggact tcggcgatta tttcatgccg ttcgccagct attcgcgcac acaccgtatg      1800
234 cccaacatcc aagaaatgta ttttctcccaa atcggcgact ccggcgttca caccgcctta      1860
235 aaaccagagc gcgcaaacac ttggcaattt ggcttcaata cctataaaaa aggattgtta      1920
236 aaacaagatg atacattagg attaaaactg gtcggctacc gcagccgcat cgacaactac      1980
237 atccacaacg tttacgggaa atggtgggat ttgaacggga atattccgag ctgggtcagc      2040
238 gacgccgggc ttgcctacac catccaacac cgcaatttca aagacaaagt acacaaacac      2100
239 ggttttgagt tggagctgaa ttacgattat gggcgtttt  tcaccaacct ttcttacgcc      2160
240 tatcaaaaaa gcacgcaacc gaccaacttc agcgatgcga gcgaatcgcc caacaatgcg      2220
241 tccaaagaag accaactcaa acaaggttat gggttgagca gggtttccgc cctgccgcga      2280
242 gattacggac gtttggaagt cggtacgcgc tggttgggca acaaactgac tttgggcggc      2340
243 gcgatgcgct atttcggcaa gagcatccgc gcgacggctg aagaacgcta tatcgacggc      2400
244 accaacgggg gaaataccag caatgtccgg caactgggca agcgttccat caaacaaacc      2460
245 gaaacccttg cccgccagcc tttgattttt gatttttacg ccgcttacga gccgaagaaa      2520
246 aaccttattt tccgcgccga agtcaaaaat ctgttcgaca ggcgttatat cgatccgctc      2580
247 gatgcgggca atgatgcggc aacgcagcgt tattacagtt cgttcgaccc gaaagacaag      2640
248 gacgaagaag taacgtgtaa tgctgataaa acgttgtgca acggcaaata cggcggcaca      2700
249 agcaaaagcg tattgaccaa ttttgcacgc ggacgcacct ttttgataac gatgagctac      2760
250 aagttttaa                                                             2769
```
252 <210> SEQ ID NO: 4
253 <211> LENGTH: 922
254 <212> TYPE: PRT
255 <213> ORGANISM: Bacteria
257 <400> SEQUENCE: 4
258 Met Arg Ser Ser Phe Arg Leu Lys Pro Ile Cys Phe Tyr Leu Met Gly
259 1   5   10   15

VERIFICATION SUMMARY   DATE: 02/27/2001
PATENT APPLICATION: US/09/762,926 TIME: 14:54:36

Input Set : A:\seqlist.txt
Output Set: N:\CRF3\02272001\I762926.raw

L:11 M:270 C: Current Application Number differs, Replaced Current Application No
L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date